

Curriculum Vitae

Mehmet M. Dalkilic

School of Informatics
Eigenmann Hall 1004
1900 East 10th Street
Bloomington, Indiana 47406
PH (812) 856-3010
FAX (812) 855-0009
dalkilic@indiana.edu

2219 Laurelwood Circle
Bloomington, Indiana 47401
PH (812) 339-6506
FAX (812) 339-6506

United States Citizen
Born: Austin, Texas

<http://www.informatics.indiana.edu/dalkilic>

Education

Ph.D. in Computer Science, Indiana University, June 2000
M.S. in Computer Science, Indiana University, 1996
B.A. in Chemistry with honors, Indiana University 1984

Experience

2004-date	Alliance for Graduate Education and the Professoriate (AGEP) faculty
2002-date	Senior Fellow Informatics Research Institute (undergraduate) Introduction to Informatics Curriculum Coordinator (graduate) Introduction to Bioinformatics Curriculum Coordinator
2001-date	Assistant Professor, School of Informatics, Indiana University Group Leader Center for Genomics & Bioinformatics Adjunct Assistant Professor Dept. of Computer Science
2000-2001	Visiting Assistant Professor, School of Informatics, Indiana University
1999-2000	Lecturer, Computer Science, Rose-Hulman Institute of Technology
1989-1998	Associate Instructor, Computer Science, Indiana University
1995-1999	Course Instructor, Elements of Discrete Structures for Computer Science, C++ I & II, Indiana University

Edited Journals / Collections

Data Mining and Bioinformatics First International Workshop, VLDB 2006, Seoul, Korea, September 11, 2006 Series: Lecture Notes in Computer Science, Vol. 4316 Sublibrary: Lecture Notes in Bioinformatics. Mehmet M. Dalkilic, Sun Kim, & Yiong Yang (Eds.) 2006, VIII, 197 p., Softcover ISBN: 978-3-540-68970-6

<http://bio.informatics.indiana.edu/VLDB06/>

Sun Kim, Zhiping Wang, and Mehmet M. Dalkilic. *iGibbs: A Motif Discovery Framework for Gibbs Sampling Algorithm*. Proteins: Structure, Function, and Bioinformatics. (2006) (in press)

<http://bio.informatics.indiana.edu/bioalgo/>

Christopher Mueller, Mehmet M. Dalkilic, and Andrew Lumsdaine. *High-Performance Direct Pairwise Comparison of Large Genomic Sequences*. Special issue of IEEE Transactions on Parallel and Distributed Systems of computational biology (2006) (in press)

Henry Kim, Arijit Sengupta, Mark Fox, and Mehmet M. Dalkilic. *A Measurement Ontology Generalizable for Emerging Domain Applications on the Semantic Web*.

A special issue of the Journal of Database Management devoted to "Defining, Eliciting and Using Data Semantics for Emerging Domains" (2006) (in press)

DoHoon Lee, Jeong-Hyeon Choi, Mehmet M. Dalkilic, and Sun Kim. *COMPAM :visualization of combining pairwise alignments for multiple genomes*. Bioinformatics (2005) 22(2): 442-444. PMID: 16269416
<http://platcom.informatics.indiana.edu/platcom/>

Mehmet M. Dalkilic. *Book review of Database Annotation in Molecular Biology: Principles and Practice*, Arthur M. Lesk (Editor). Briefings in Bioinformatics 6(3) (2005) (invited)

Using History in an Introductory Informatics Course. Mehmet Dalkilic. Using History to Teach Computer Science and Related Disciplines. Atsushi Akera and William Aspray Eds. (2004) pp. 147-157 (invited)

Mehmet M. Dalkilic, Edward Robertson, and Dirk Van Gucht. {Book Chapter} *CE: The Classifier-Estimator Framework for Data Mining* in "Data Mining and Reverse Engineering," Spaccapietra, Stefano; Maryanski, Fred (Eds.) 1998, 520 p., Hardcover ISBN: 0-412-82250-4.

Refereed Conferences and Workshops

James Costello, Jade Buchanan-Carter, Mehmet Dalkilic and Justen Andrews
Integrating Drosophila Data to Discover Disease-Related Protein Interactions in Human, 2007 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (2007) (to appear)
<http://www.monkey.informatics.indiana.edu/disease/>

Sukamol Srikwan, Markus Jakobsson, Andrew Albrecht, Mehmet Dalkilic. *Trust Establishment in Data Sharing: An Incentive Model for Biodiversity Information Systems*. The International Workshop on Trusted Collaboration (TrustCol-2006) (to appear)

Patwardhan R, Tang H, Kim S and Dalkilic M. *An Approximate De Bruijn Graph Approach to Multiple Local Alignment and Motif Discovery in Protein Sequences*, 1st VLDB Workshop on Data Mining in Bioinformatics, LNCS (to appear)
<http://biokdd.informatics.indiana.edu/rpatward/debruijn/project.html>

Mehmet M. Dalkilic, Wyatt T. Clark, James C. Costello, and Predrag Radivojac.
Using Compression to Identify Classes of Inauthentic Texts. SIAM Conference on Data Mining (2006) (in press)
<http://www.inauthentic.org>

Christopher Mueller, Mehmet Dalkilic, and Andrew Lumsdaine. *High-Performance Direct Pairwise Comparison of Large Genomic Sequences*. In Proceedings of the Fourth IEEE International Workshop on High Performance Computational Biology (HiCOMB 2005), (2005) p. 199

Mehmet M. Dalkilic and Arijit Sengupta. *Design and evaluation of CATPA: curation and alignment tool for protein analysis*. ACM SAC '05: Proceedings of the 2005 ACM symposium on Applied computing, Track: Bioinformatics, Santa Fe, New Mexico (2005) pp. 190-194, ISBN: 1-58113-964-0
<http://www.catpa.org>

Mehmet Dalkilic. *An Architecture and Application for Integrating Curation Data at the Residue Level for Proteins*. Data Integration in the Life Sciences, Second International Workshop. Lecture Notes In Computer Science; Vol. 3615. San Diego, CA, July 20-22 (2005) pp. 335-338, ISBN: 3-540-27967-9

Mehmet Dalkilic and Arijit Sengupta. *Circle: Design and Implementation of a Classifier Based on Circuit Minimization*. ACM Symposium on Applied Computing, Special Track on Data Mining (Short paper) Santa Fe, NM (2005) pp. 547-548, ISBN: 1-58113-964-0
<http://www.wright.edu/~arijit.sengupta/circle/>

Transforming Biology into an Information Science. Mehmet Dalkilic, Sun Kim, Predrag Radivojac, and Haixu Tang. First Inaugural i-Conference 2005. The Pennsylvania State University, PE (2005).

Mehmet Dalkilic and Arijit Sengupta. *Semantic Thumbnails: Summarizing XML Documents for the Semantic Web*. In Proceedings, XML 2004 Conference, Washington, DC. November 15-19, 2004. (Available directly from XML 2004 Online Proceedings).

Mehmet Dalkilic and James Costello. *BioKnOT: Biological Knowledge through Ontologies and TFIDF* in SIGIR Bioinformatics Workshop in conjunction with SIGIR Sheffield, UK (2004)

Zhiping Wang, Mehmet Dalkilic and Sun Kim. *Guiding motif discovery by iterative pattern refinement*. In Proceedings of the 2004 ACM symposium on Applied computing Bioinformatics Session, Nicosia, Cyprus (2004) pp. 162-166

Arijit Sengupta, Mehmet Dalkilic and James Costello. *Semantic Thumbnails - a Novel Method for Summarizing Document Collections* In Proceedings, 22nd ACM Annual International Conference on Design of Communication (SIGDOC 2004), Memphis, TN. USA. (2004) pp. 45-51, ISBN: 1-58113-809-1

Irfan Gunduz, Sihui Zhao, Mehmet Dalkilic and Sun Kim. *Motif Discovery from Large Number of Sequences: A Case Study with Disease Resistance Genes in Arabidopsis thaliana*. Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences, METMBS '03, Las Vegas, Nevada, USA, June 23 - 26, (2003) pp 29-34, ISBN: 1-932415-04-1. Faramarz Valafar and Homayoun Valafar (Eds.)

Chris M. Giannella, Mehmet M. Dalkilic, Dennis P. Groth, Edward L. Robertson. *Improving Query Evaluation with Approximate Functional Dependency Based Decompositions*. Lecture Notes In Computer Science; Vol. 2405. Proceedings of the 19th British National Conference on Databases: Advances in Databases BNCOD 19, Sheffield, UK, July 17-19 (2002) pp 26-41, ISBN: 3-540-43905-6

Mehmet Dalkilic, Edward Robertson. *Information Dependencies*. Proceedings Of the 19th ACM SIGACT-SIGMOD-SIGART Symposium on the Principles of Database Systems (extended abstract) (2000) pp. 245-253.

Mehmet Dalkilic, Edward Robertson, Dirk Van Gucht. *CE: The Classifier-Estimator Framework for Data Mining*. Proceedings of IFIP TC2/WG2.6 Seventh Conference on Database Semantics (DS-7) Leysin, Switzerland. October 7-10, (1997) pp 89-104.

Grants

Integration of functional relationships among genes in Drosophila. \$100K (2006) Justen Andrews, John Colbourne, Mehmet Dalkilic

<http://www.indigene.org>

NIH Exploratory Center for Chemicalformatics. \$360K (Significant Contributor)

Microsoft eScience \$15K PI (2006)

<http://www.monkey.informatics.indiana.edu>

Multidisciplinary Ventures and Seminars Fund. \$6K PI, Co-PI Claudia Johnson (Geology) and Co-PI Erika Elswick (Geology) (2005-2006)

IBM Life Sciences/UITs Start-up for CATPA \$70K. PI (2002)

IBM Shared University Research grant. "Grid and Data Intensive Computing in the Life Sciences." \$500K PIs: Craig A. Stewart, Michael A. McRobbie, Ora Pescovitz, Beth Plale, Howard Edenberg, Mehmet Dalkilic, John C. Huffman, Peter Cherbas, Sun Kim, Vadim Moskvina, Mary Papakhian, Anurag Shankar, Richard Repasky, David Hart, Donald Steward, George Turner (2001)

Data Mining and Information Dependencies, National Science Foundation (2001) IIS-0082401, \$420K, Co-PI with PI Edward Robertson and Co-PI Dirk Van Gucht.

Curriculum

Undergraduate

I101 Introduction to Informatics*

18 Lecture Sections/ ~70 Laboratories /~ 1000 Students

I201 Mathematical Foundations

One Lecture Section/ 1 Laboratory / ~ 80 Students

I400 e-Tools

One Lecture / 1 Laboratory / ~15 Students

I400 Data Mining

One Lecture / 1 Laboratory / ~ 35

Graduate

L519 Introduction to Bioinformatics and Computational Biology*

3 Lecture Sections / 3 Laboratories / ~ 50 Students

L529 Bioinformatics and Computational Biology

1 Lecture Section / 1 Laboratory / ~15 Students

Capstone Advisor (responsible for managing year-long Master's projects)

<http://www.informatics.indiana.edu/academics/capstone.asp>

*Curriculum Coordinator—responsible for assessing, designing, and implementing course content. Capstone became formal class 2005. Shared responsibility with Sun Kim, Haixu Tang; Formally responsible for class twice. More than 1200 students taught since 2000.

Highschool & Junior High

Jim Holland Summer Enrichment Program in Biology, Introduction to Bioinformatics Science for under-represented groups. Indiana University, Bloomington, IN

Informatics Summer IT Camp: Bioinformatics School of Informatics, (2004-

2006) Indiana University, Bloomington, IN (Junior High)

Talks

Network Properties of a Gene Network Built from Drosophila melanogaster Data. International Conference on Network Science, Bloomington, IN. May 22-25 (2006)

Design and evaluation of CATPA: curation and alignment tool for protein analysis. ACM SAC '05: Proceedings of the 2005 ACM symposium on Applied computing, Track: Bioinformatics, Santa Fe, New Mexico. March 13-17 (2005)

Guiding motif discovery by iterative pattern refinement. In Proceedings of the 2004 ACM symposium on Applied computing Bioinformatics Session, Nicosia, Cyprus. March 14-17 (2004)

Semantic Thumbnails: Summarizing XML Documents for the Semantic Web. XML 2004 Conference, Washington, DC. November 15-19 (2004)

Semantic Thumbnails - a Novel Method for Summarizing Document Collections. 22nd ACM Annual International Conference on Design of Communication, Memphis, TN. USA. Oct 10-14 (2004)

How to build a curriculum from scratch: I101 Introduction to Informatics Mehmet Dalkilic International Society for the Scholarship of Teaching and Learning Inaugural Conference. Bloomington, IN. Oct 22-24 (2004)

CE: The Classifier-Estimator Framework for Data Mining. Seventh Conference on Database Semantics (DS-7) Leysin, Switzerland. October 7-10 (1997)

Invited Talks

Paleoinformatics in the PASTT: Information technology is not just for the future anymore, Mehmet Dalkilic, Claudia C. Johnson, Erika R. Elswick, Scott Beason. Microsoft eScience Workshop at Johns Hopkins (2006) (invited poster)

From Combinatorics to Statistics: Motif Discovery. Algorithmic Biology: Algorithmic Techniques in Computational Biology, National University of Singapore, Singapore. June 1-July 31 (2006)

Tutorial on Systems biology. Algorithmic Biology: Algorithmic Techniques in Computational Biology. National University of Singapore, Singapore, July 8-14 (2006)

An architecture and application for integrating curation data at the residue level for proteins. Mehmet Dalkilic. 2nd International Workshop on Data Integration in the Life Sciences. San Diego, CA. July 20-22 (2005)

Integration and Visualization in Bioinformatics. Mehmet Dalkilic. eScience Workshop, Microsoft. Redmond, Seattle Oct 6-8 (2005)

A logic-theoretic classifier called Circle. IEEE 8th International Conference on Control Automation, Robotics and Vision, Kunming, China (ICARCV 2004)*

The Management of Biological Information. Executive IT Life Sciences Forum. Washington, DC. Nov. 30 (2003)

Technological Waterloo. The Johnson Center for Entrepreneurship and Innovation (The Kelley School of Business). Indianapolis, IN. Feb. 20 (2002)

Technical Reports/Posters

Mehmet Dalkilic and Arijit Sengupta. (Poster) *Circle: Design and Implementation of a Classifier Based on Circuit Minimization*. ACM Symposium on Applied Computing, Special Track on Datamining Santa Fe, NM (2005)

Mehmet Dalkilic and Mike Groomer. *On Comparing and Measuring Continuous Audit Streams*. Accepted for Presentation at the Eighth World Continuous Auditing and Reporting Symposium, Newark, NJ (2004)

Mehmet Dalkilic and Arijit Sengupta. *Circle: A logic-theoretic classifier for sequence analysis*. (Poster) First Midwest Database Symposium, Chicago, Ill April (2004)

Design & Implementation of Reflective SQL, Mehmet Dalkilic, Manoj Jain, Anurag Mendhekar (TR 451)

Awards and Honors

Student Choice Award (2004-2005)

School of Informatics Teaching Excellence Award (2003-4)

School of Informatics Teaching Excellence Award (2001-2)

Favorite Professor of the Year, Freshmen Honor Society Alpha Xi Omega (2001-2)

Professional Activities

VLDB 2006 Workshop: Data Mining in Bioinformatics co-chair (Mehmet M. Dalkilic, Sun Kim, Jiong Yang) (Seoul, Korea)

Bio-Viz Symposium '06 Visualization in Bioinformatics (Sydney, Australia) co-chair (Mehmet M. Dalkilic and Dennis Groth)

Indiana University Faculty Learning Community Participant (Summer 2005)

Outreach Program: Informatics IT Summer Camp . This includes providing hands-on IT activities that expose students to the field of Science, Computer Science, & Bioinformatics. (2004-2006)

Professional Activities (continued)

Program Committees/Reviewer: ACM (many), PSB, ISMB, VLDB, ISMB, IEEE (many), Journal of History and Philosophy of Science, *etc.* 2000-to date

NSF Panel Reviewer (2003)

Member ACM, attendee to SIGMOD/PODS 1992-to date

Member IEEE, 2000-to date

Microsoft Faculty Summit (2005-2006)

AGEP Faculty (Mentor Under-represented students for graduate school)